A F $

E f w

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: :

Ana Belen Benitez et al. : Attorney Ref.: 2000-0025

:

Serial No.: 09/495,175 : Confirmation No.: 4490

:

Filed: February 1, 2000 : Art Unit: 2142

:

FOR: MULTIMEDIA INTEGRATION : Examiner: Beatriz Prieto
DESCRIPTION SCHEME, METHOD :
AND SYSTEM FOR MPEG-7 :

APPEAL BRIEF

MAIL STOP: Appeal Brief Patents
Commissioner for Patents
Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sirs:

This is an Appeal from an Office Action dated January 23, 2004, finally rejecting

each of the pending claims, 25 - 53. The Notice of Appeal was filed on June 23, 2004. A

Petition for the Extension of Time is being filed in conjunction with this Appeal Brief,

extending the filing period until September 23, 2004.

1. **REAL PARTY IN INTEREST**

The real parties in interest in this matter are AT&T Corp. and Columbia University.

2. **RELATED APPEALS AND INTERFERENCES**

There are no related appeals or interferences.

3. **STATUS OF THE CLAIMS**

This appeal is an appeal from the rejection of pending claims 25 - 53.

09/27/2004 DEHMANU1 00000039 09495175

01 FC:1402   330.00 OP

1

4. STATUS OF AMENDMENTS

Appellants make one minor amendment to claim 25 in this appeal. As this amendment is governed by 37 C.F.R. Section 1.116, Appellants note that this amendment merely clarifies that the entity relation graph descriptions for that at least one of the multimedia types are non-hierarchical. Support for this claim is found on in the specification on page 18. Appellants submit this amendment to place the claims in better form for consideration on appeal.

5. SUMMARY OF THE INVENTION

The present inventions relate to the MPEG-7 standard for searching and organizing multimedia data. The invention provides a method for generating description records from multimedia content. The invention allows for the integration of the descriptions of categories of audiovisual content to allow the content to be searched or browsed with ease in digital libraries. The method embodiment of the invention comprises identifying multimedia types in multimedia content, extracting multimedia objects to generate multimedia object descriptions from the multimedia content for at least one of the multimedia types, generating, from the multimedia object descriptions, non-hierarchical entity relation graph descriptions for at least one of the multimedia types and integrating the multimedia object descriptions and the entity relation graph descriptions to generate at least one description record to represent content embedded within the multimedia content.

Another aspect of the invention is a method of generating description records from multimedia content. The method comprises identifying multimedia types in multimedia content, extracting multimedia objects to generate multimedia object descriptions from the multimedia content for at least one of the multimedia types, generating, from the multimedia object descriptions, multimedia object hierarchy descriptions by object hierarchy construction and extraction processing for at least one of the multimedia types and integrating the multimedia object descriptions and the multimedia object hierarchy descriptions to generate at least one description record to represent content embedded within the multimedia content.

6.   **ISSUES**

(1)   Whether the Examiner's interpretation of the teachings of Suzan et al. and the interpretation of the claims are correct.

(2)   Whether one of skill in the art would find it obvious to combine Suzan et al. with Burgman et al.

7.   **GROUPING OF CLAIMS**

The claims may be grouped as follows.

Claims 25 - 52 relate one method embodiment of the invention; and

Claim 53 relates to another method embodiment of the invention.

A separate basis for patentability exists for each group of claims. However, except to the extent otherwise indicated below, the respective groups of claims stand or fall together for purposes of this appeal.

8.   **ARGUMENT**

MPEG-7, like all standards, progressively changed as the committees discussed and agreed on each requirement. Therefore, the statements herein do not apply to the final adopted version of MPEG-7, which has different requirements for the Description Scheme (DS) that those summarized in Suzan et al. The arguments herein clearly only apply to the specific Suzan et al. article and how their summary of the requirements for MPEG-7 at the time.

(A)   **The Examiner incorrectly interprets of the teachings of Suzan et al. and the claims**

The Examiner rejects claims 25 - 53 under Section 103 as being unpatentable over Suzan et al. MPEG-7 Standardization Activities ("Suzan et al.") in view of U.S. Patent No. 6,564,263 to Bergman et al. ("Bergman et al."). Appellants have traversed this rejection and

have provided a number of arguments regarding why the Examiner is incorrectly interpreting the teachings of both the prior art and the claims.

Independent claim 25 recites a method for generating description records from multimedia content. The third step in the claim requires generating from the multimedia object descriptions, non-hierarchical entity relation graph descriptions for at least one of the multimedia types. Appellants note that Suzan et al. as a reference is reporting on the current status of the development of the MPEG-7 standard. Suzan et al. set forth the details of the MPEG-7 protocol and the technical patterns necessary to comply with the protocol. In other words, Suzan et al. is teaching one of skill in the art about what must be done to comply with MPEG-7 at is existed at the time. They note on page 520 that "As it is apparent from above, MPEG-7 is still in its formation process. Discussions are still ongoing in an effort to determine the approach to the best possible standard." Under Section 3 of Suzan et al., they set forth "Requirements of MPEG-7" and list features for audio and visual requirements that must be complied with to match the standard. In this context, Suzan et al. notably would not suggest to one of skill in the art to seek technical features outside the current proposed protocol requirements. The reason for this is that any such excursion would eliminate compliance with MPEG-7 as it then existed.

The Examiner in the Office Action dated January 23, 2004 relies on the "Feature hierarchy" section of Suzan et al. as matching disclosure for the entity relation graph descriptions of claim 25. The feature hierarchy section states: *"Feature hierarchy:* MPEG-7 shall support the hierarchical representation of different features in order that queries may be processed more efficiently in successive layers." (emphasis in original). Appellants previously explained that the present specification supported an interpretation of the entity relation graphs as non-hierarchical. Appellants also argued that the Examiner improperly interpreted Suzan et al. and claim 25 to equate the reference with the claim limitations. The present amendment to claim 25 makes clear that the entity relation graphs are non-

4

hierarchical. Therefore, Suzan et al. actually teach away from the present invention because they require a hierarchical representation and claim 25 recites a non-hierarchical entity relation graph.

With regards to whether one of skill in the art would read Suzan et al. and find any suggestion to use a hierarchical approach, Suzan et al. make clear that there are no exceptions to the hierarchical requirement. At the time of the Suzan et al. reference, the MPEG-7 standard was required to support the hierarchical representation of different features so that queries may be processed more efficiently. Therefore, Appellants submit that Suzan et al. actually teach away from the non-hierarchical approach recited in claim 25. For the foregoing reasons, Appellants submit that claim 25 is patentable over the cited combination of references.

Claims 26 - 52 each depend from claim 25 and recite further limitations therefrom. Each of these claims is therefore allowable inasmuch as the parent claim is allowable.

(B)   One of skill in the art would not find it obvious to combine Suzan et al. with Burgman et al.

In addition to the interpretation of Suzan et al. and claim 25 issue above, Appellants submit that there is no reason or suggestion to combine Suzan et al. with Burgman et al. The requirements on the Examiner for establishing the *prima facie* case of obviousness are set forth in the June, 2004 Response.

The Examiner conceded that Suzan et al. do not teach the limitation of integrating the multimedia descriptions and the entity relation graph descriptions to generate a description record to represent content embedded within the multimedia content. Amended claim 25 expressly makes the entity relation graphs non-hierarchical. The Examiner then asserted that Bergman et al. provide this missing limitation and that it would be obvious to combine Suzan et al. with Bergman et al. because one of skill in the art would be motivated "to generate

entity relation description based on the multimedia object descriptions for multiple

multimedia content types including a composite multimedia object description that represents

content embedded within the multimedia content, as taught by Bergman."

Petitions submit that there is no suggestion or motivation to combine Bergman et al.

with Suzan et al. for several reasons. First, Bergman et al. was filed as a provisional in

December, 1998. Suzan et al. was published October 1998 as a conference proceeding

reporting on the then-current proposal for MPEG-7. The purpose of Bergman et al. was to

introduce a multimedia content description framework that differed from the MPEG-7

submissions. In this regard, Bergman et al. state:

> In one multimedia application relating to the Motion Picture Experts Group (MPEG)
> data compression standard, for example, a main difference between the InfoPyramid
> approach of the present invention and other conventional schemes proposed for
> MPEG-7, is the virtually complete elimination of the dichotomy between data and
> metadata. Col. 9, lines 39 - 44.

As can be seen from this cited paragraph, Appellants submit that Bergman et al.

clearly distance their InfoPyramid approach from the current proposals for MPEG-7. The

dates of the Bergman et al. application and the Suzan et al. reference is instructive in that they

are almost exactly concurrent in time. Note further that Suzan et al. is based on proposals for

MPEG-7 standards submitted in March, 1998. Given the timing of these prior art documents,

the current proposals for MPEG-7 referenced by Bergman et al. may be represented by the

MPEG-7 summary in Suzan et al.

Therefore, given the fact that Bergman et al. explicitly states that their invention

differs from the convention MPEG-7 approach, Appellants submit that Bergman et al.

actually teach away from any combination with Suzan et al.'s articulation of the current state

of MPEG-7. Appellants respectfully submit that Bergman et al. would certainly not suggest

to one of skill in the art to blend their teachings with Suzan et al., which sets forth MPEG-7

as it then existed. This fact is further bolstered by the well-known experience in developing

standards where various contributors strive to have their proposed technology adopted into

the standard. In this regard, Bergman et al. desired that their InfoPyramid content description framework would be adopted over the then-current MPEG-7 proposal. Therefore, one of skill in the art would not be motivated to combine or blend these divergent suggestions for adoption into a single standard.

The obviousness analysis requires one to place him/herself in the time the "invention was made" 35 U.S.C. Section 103(a). The Examiner must determine, at the time the invention was made, whether it would be obvious to one of skill in the art to combine the references. The present application was filed in February 2000. At this time, the MPEG-7 had not yet been finalized. (See MPEG-7 schedule on page 517, Suzan et al.). Therefore, one of skill in the art would know that the standards body would still be considering a working draft of MPEG-7 in February 2000. This information is provided to give the Board the context of what a person of skill in the art would understand Bergman et al.'s statement about being different from MPEG-7 proposals to mean in February 2000.

For these reasons, Appellants respectfully submit that one of skill in the art would not have found motivation to modify Suzan et al. with Bergman et al. or to blend the teachings of these two references.

Even if Suzan et al. were combined with Bergman et al., Appellants submit that the combination still fails to teach each element of Claim 25. First, Appellants submit that Bergman et al. clearly do teach a hierarchical approach with their InfoPyramid 302 and associated disclosure. FIG. 4 and FIG. 9 of Bergman et al. show the pyramid and how the relationship between elements in the pyramid can only be described as hierarchical. For example, within the pyramid, there is a highest level of fidelity or resolution along the base of the pyramid and the lowest level of fidelity or resolution along the top of the pyramid. Varying parameter degrees are in between. Col. 7, lines 14 - 25. Therefore, Bergman et al. teach that the pyramid organizes data according to a hierarchy of highest values to lowest values within the pyramid. While the Infopyramid of Bergman et al. differs from the feature

hierarchy mentioned in Suzan et al., both these approaches do have a hierarchical component. Therefore, even if they were combined, Suzan et al. and Bergman et al. fail to teach a non-hierarchical entity graph as recited in claim 25.

If the Examiner were to assert that Bergman et al. teach a non-hierarchical approach, and thus their disclosure matches the entity relation graph of claim 25, then there could not be any reason to combine its teachings with Suzan et al. The reason for the lack of motivation to combine is found in the fact that Suzan et al. (page 519) require for compatibility with MPEG-7 a feature hierarchical representation of different features. As set forth above, Bergman et al. explicitly state that they are proposing an approach that differs from MPEG-7 proposals. Therefore, one of skill in the art would not look to a non-hierarchical approach where MPEG-7 at that time required a hierarchical solution.

A further reason that prevents one of skill in the art from having motivation to combine Suzan et al. with Bergman et al. also relates to the fact that Suzan et al. is a summary of the current status of MPEG-7. The Examiner states that it would be obvious to include Bergman et al.'s teachings regarding aggregating, incorporating or combining the multimedia object descriptions and their entity graph relationship graph descriptions to generate the description record to represent content embedded within the multimedia content. A problem with this reasoning is that it requires one of skill in the art to incorporate Bergman et al.'s scheme into Suzan et al., which would then render the teachings of Suzan et al. outside the MPEG-7 proposed protocol. For example, the "description record" in Suzan et al. is described on page 518 in the following paragraph: *"Description:* A description is the entity describing the data and consisting of a DS [Description Scheme] and instantiation of the corresponding descriptors." The DS and descriptors are also set forth on page 518.

If Bergman et al.'s approach to creating a content description framework were employed, it would differ dramatically from the then-current MPEG-7 approach of Suzan et al. As mentioned above, Bergman et al. specifically state that their approach is different from

MPEG-7. Therefore, if Bergman et al. teach what the Examiner asserts that they teach, then their integrated description record that represents content embedded within multimedia content clearly differs from the then-current MPEG-7 as set forth in Suzan et al. For this reason, one of skill in the art would not find motivation to go outside the MPEG-7 protocol, and change the requirements of the protocol with subject matter from Bergman et al. As set forth in the MPEP, if the blending of two references requires a change in the principle of operation of a reference, then there can be no motivation to combine.

Therefore, Appellants submit that claim 25 is patentable in that there is no reason or suggestion to combine Suzan et al. with Bergman et al. Furthermore, claims 26 - 52 each depend from claim 25 and recite further limitations therefrom. Accordingly, Appellants submit that these claims are patentable as well.

The Examiner also combined Suzan et al. with Bergman et al. to reject claim 53. Given the explanation above, Appellants submit that claim 53 is patentable since there is no motivation or suggestion to combine these two references.

SUMMARY

In view of the above, Appellants submit that all claims on appeal distinguish over the art and respectfully request that the Examiner's rejections of these claims be reversed. Appellants, therefore, respectfully move this Board to reverse the Examiner's decision rejecting claims 25 - 53.

Respectfully submitted,

Date: September 23, 2004

By: _Thomas M. Isaacson_

Thomas M. Isaacson
Attorney for Appellants
Reg. No. 44,166
Phone: 410-414-3056
Fax No.: 410-510-1433

Correspondence Address:
Samuel H. Dworetsky
AT&T Corp.
Room 2A-207
One AT&T Way
Bedminster, NJ 07921

APPENDIX

(Brief of Appellants Benitez et al.
U.S. Patent Application Serial No. 09/495,175)

<u>CLAIMS ON APPEAL</u>

25. (currently amended) A method for generating description records from multimedia content, comprising:

identifying multimedia types in multimedia content;

extracting multimedia objects to generate multimedia object descriptions from the multimedia content for at least one of the multimedia types;

generating, from the multimedia object descriptions, <u>non-hierarchical</u> entity relation graph descriptions for at least one of the multimedia types; and

integrating the multimedia object descriptions and the entity relation graph descriptions to generate at least one description record to represent content embedded within the multimedia content.

26. (Previously Presented) The method of claim 25, further comprising generating, from the multimedia object descriptions, multimedia object hierarchy descriptions by object hierarchy construction and extraction processing, for at least one of the multimedia types.

27. (Previously Presented) The method of claim 25, wherein the multimedia types include at least one of image, audio, video, synthetic, and text.

28. (Previously Presented) The method of claim 25, wherein the extracting of multimedia objects further comprises:

segmenting each multimedia content into segments including content from at least one of the multimedia types for the multimedia content; and

generating at least one feature description for at least one of the segments by feature extraction and annotation;

wherein the generated multimedia object descriptions comprise the at least one feature description for the at least one segment.

29. (Previously Presented) The method of claim 28, wherein the segments are selected from the group consisting of local segments and global segments.

30. (Previously Presented) The method of claim 28, further comprising selecting the at least one feature description from the group consisting of media, semantic and temporal features.

31. (Previously Presented) The method of claim 30, wherein the media features are further defined by at least one feature description selected from the group consisting of data location, scalable representation and modality transcoding.

32. (Previously Presented) The method of claim 30, wherein the semantic features are further defined by at least one feature description selected from the group consisting of keywords, who, what object, what action, why, when, where and text annotation.

33. (Previously Presented) The method of claim 30, wherein the temporal features are further defined by at least one feature description consisting of duration.

34. (Previously Presented) The method of claim 28, wherein the extracting of multimedia objects further comprises:

　　　　generating media object descriptions from the multimedia segment for one of the multimedia types by media object extraction processing;

　　　　generating media object hierarchy descriptions from the generated media object descriptions by object hierarchy construction and extraction processing; and

　　　　generating media entity relation graph descriptions from the generated media object descriptions by entity relation graph generation processing.

35. (Previously Presented) The method of claim 34, wherein generating media object descriptions further comprises:

　　　　segmenting the content of each multimedia type in the multimedia object into segments within the multimedia object by media segmentation processing; and

　　　　generating at least one feature description for at least one of the segments by feature extraction and annotation;

　　　　wherein the generated media object descriptions comprise the at least one feature description for the at least one of the segments.

36. (Previously Presented) The method of claim 35, further comprising the step of selecting the at least one feature description from the group consisting of media, semantic and temporal.

37. (Previously Presented) The method of claim 35, wherein generating media object hierarchy descriptions generates media object hierarchy descriptions of the media object descriptions based on media feature relationships of media objects represented by the media object descriptions.

38. (Previously Presented) The method of claim 35, wherein generating media object hierarchy descriptions generates semantic object hierarchy descriptions of the media object descriptions based on semantic feature relationships of media objects represented by the media object descriptions.

39. (Previously Presented) The method of claim 35, wherein generating media object hierarchy descriptions generates temporal object hierarchy descriptions of the media object descriptions based on temporal features relationships of media objects represented by the media object descriptions.

40. (Previously Presented) The method of claim 35, wherein generating media object hierarchy descriptions generates media object hierarchy descriptions of the media object descriptions based on relationships of media objects represented by the media object descriptions, and wherein the relationships are selected from the group consisting of media feature relationships, semantic feature relationships, temporal feature relationships, and spatial feature relationships.

41. (Previously Presented) The method of claim 35, wherein generating media entity relation graph descriptions generates entity relation graph descriptions of the media object descriptions based on relationship of the media objects represented by the media object descriptions, wherein the relationships are selected from the group consisting of media feature relationships, semantic feature relationships, temporal feature relationships and spatial feature relationships.

42. (Previously Presented) The method of claim 26, wherein generating multimedia object hierarchy descriptions generates multimedia object hierarchy descriptions of the multimedia object descriptions based on media feature relationships of multimedia objects represented by the multimedia object descriptions.

43. (Previously Presented) The method of claim 26, wherein generating multimedia object hierarchy descriptions generates semantic object hierarchy descriptions of the multimedia object descriptions based on semantic feature relationships of multimedia objects represented by the multimedia object descriptions.

44. (Previously Presented) The method of claim 26, wherein generating multimedia object hierarchy descriptions generates temporal object hierarchy descriptions of the multimedia object descriptions based on temporal feature relationships of multimedia objects represented by the multimedia object descriptions.

45. (Previously Presented) The method of claim 26, wherein generating multimedia object hierarchy descriptions generates multimedia object hierarchy descriptions of the multimedia object descriptions based on relationships of multimedia objects represented by the multimedia object descriptions, wherein the relationships are selected from the group consisting of media feature relationships, semantic feature relationships, temporal feature relationships and spatial feature relationships.

46. (Previously Presented) The method of claim 25, wherein generating entity relation graph descriptions generates the entity relation graph descriptions of the multimedia object descriptions based on relationships of multimedia objects represented by the multimedia object descriptions, wherein the relationships are selected from the group consisting of media feature relationships, semantic feature relationships, temporal feature relationships and spatial feature relationships.

47. (Previously Presented) The method of claim 25, further comprising receiving and encoding the multimedia object descriptions into encoded description information, and storing the encoded description information as the at least one description record.

48. (Previously Presented) The method of claim 26, further comprising combining the multimedia object description, the multimedia object hierarchy descriptions, and the entity relation graph description to form a multimedia description, and receiving and encoding the multimedia description into encoded description information, and storing the encoded description information as the at least one description record.

49. (Previously Presented) The method of claim 47, wherein the encoding comprises binary encoding.

50. (Previously Presented) The method of claim 48, wherein the encoding comprises binary encoding.

51. (Previously Presented) The method of claim 47, wherein the encoding comprises the extensible Markup Language (XML) encoding.

52. (Previously Presented) The method of claim 47, wherein the encoding comprises the extensible Markup Language (XML) encoding.

53. (Previously Presented) A method of generating description records from multimedia content, comprising:

identifying multimedia types in multimedia content;

extracting multimedia objects to generate multimedia object descriptions from the multimedia content for at least one of the multimedia types;

generating, from the multimedia object descriptions, multimedia object hierarchy descriptions by object hierarchy construction and extraction processing, for at least one of the multimedia types; and

integrating the multimedia object descriptions and the multimedia object hierarchy descriptions to generate at least one description record to represent content embedded within the multimedia content.